Introduction to Ontology

Lecture 01

1

Representing Knowledge

- In computing and intelligent software development, an ontology is a manipulable object.
- Ontologies can be presented in multiple formats, depending on the intended user.
- Fundamentally, an ontology is a logic-based representation that a computer can process.

2

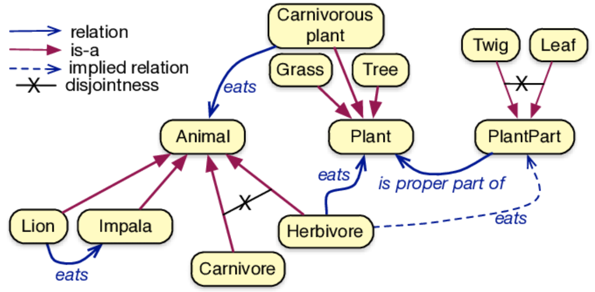
Representing Knowledge

Example: African Wildlife Ontology (AWO)

- AWO contains knowledge about wildlife, for example:
  - Giraffes eat leaves and twigs.
  - Giraffes are herbivores.
  - Herbivores are animals.

3

Representing Knowledge



4

Representing Knowledge

First-order Predicate Logic

- A mathematician may prefer to represent such knowledge with first order predicate logic.

$$\forall x(Lion(x) \rightarrow \forall y(eats(x,y) \rightarrow Herbivore(y)) \wedge \exists z(eats(x,z) \wedge Impala(z)))$$

- “All lions eat herbivores, and they also eat some impalas".
- This axiom may be one of the axioms in the ontology.

5

5

Representing Knowledge

Description Logic language

- Knowledge can be represented in logics other than plain vanilla first order logic.
- Description Logic (DL) is one such formalism for representing knowledge.
- For example, the given knowledge can be expressed as:

$$\text{Lion} \sqsubseteq \forall \text{eats.Herbivore} \sqcap \exists \text{eats.Impala}$$

6

6

Representing Knowledge

Natural Language Sentences

- Domain experts generally prefer user-friendly representations of formal expressions.
- One approach is the automatic generation of pseudo-natural language sentences.

Each lion eats only herbivore and eats some Impala

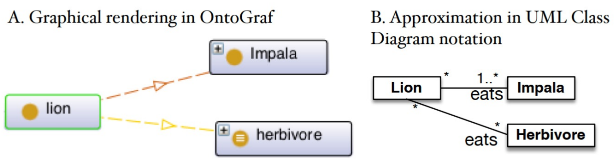
7

7

Representing Knowledge

Graphical Language

- Use of graphical language can represent knowledge with varying degrees of precision.



8

8

Representing Knowledge

- An ontology is an engineering artefact designed to represent knowledge.
- It must have a machine-processable format to enable computational use.
- The format should faithfully adhere to formal logic.
- Above discussed knowledge representations are not easily processable by computers.

9

9

Representing Knowledge

- Ontologies can be serialized into text files to enable computer processing.
- Most widely used ontology serialization language is the Web Ontology Language (OWL).
- OWL requires a specific format called **RDF/XML** for machine-processable representations.

10

10

Representing Knowledge

```
<owl:Class rdf:about="&AWO;lion">
   <rdfs:subClassOf rdf:resource="&AWO;animal"/>
   <rdfs:subClassOf>
      <owl:Restriction>
        <owl:onProperty rdf:resource="&AWO;eats"/>
        <owl:someValuesFrom rdf:resource="&AWO.owl;Impala"/>
      </owl:Restriction>
   </rdfs:subClassOf>
   <rdfs:subClassOf>
      <owl:Restriction>
        <owl:onProperty rdf:resource="&AWO;eats"/>
        <owl:allValuesFrom rdf:resource="&AWO;herbivore"/>
      </owl:Restriction>
   </rdfs:subClassOf>
   <rdfs:comment>Lions are animals that eat only herbivores.</rdfs:comment>
</owl:Class>
```

11

11

Representing Knowledge

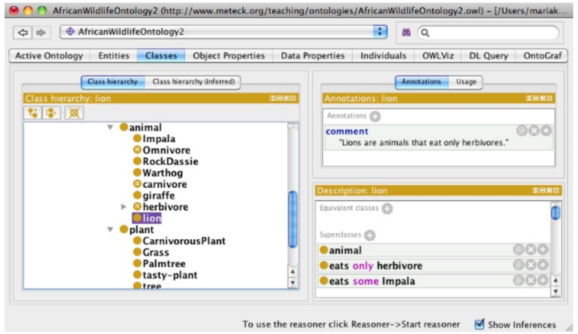
- Writing an ontology directly in RDF/XML format is generally not required.
- Computer scientists may design tools to process or modify machine-readable ontology files.
- Ontology authoring can be carried out using **Ontology Development Environments** (ODEs).
- These environments provide representations of ontologies in graphical, textual, or logical views.

12

12

Representing Knowledge

Example: Prot´eg´e



13

13

What is an Ontology?

Ontologies Vs Conceptual Data Models.

Conceptual Data Models	Ontologies
Provide application-specific yet implementation-independent representations of data.	Offer application-independent representations of a specific subject domain.
Focus on the data requirements of a particular application.	Are reusable across multiple applications regardless of context.
Usually informal, represented through diagrams (boxes & lines)	Typically formalized using logic-based languages

14

14

What is an Ontology?

Ontologies Vs Relational Databases

- RDBMSs differ from ontologies in their approach to knowledge representation.
- Ontologies explicitly represent knowledge, incorporating rules within the knowledge base.
- They employ automated reasoning to infer implicit knowledge and identify inconsistencies.
- Ontologies typically operate under the **Open World Assumption (OWA)**, unlike RDBMSs.

15

15

What is an Ontology?

- Above comparison provided is informal and vague, offering only a general sense of what an ontology might or might not be, without delivering a precise definition.
- In 2007, the Ontolog Communiqué addressed the issue of defining ontology by presenting a diagram illustrating various elements related to the concept.

16

16

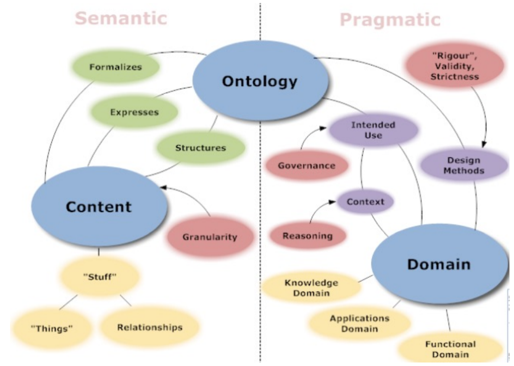
What is an Ontology?

- This diagram was intended to serve as a "**Template for discourse**" about ontologies, accompanied by both brief and detailed explanations of the elements represented.
- Framework distinguishes between two aspects:
- **Semantic aspect** – concerning the meaning conveyed by the ontology.
- **Pragmatic aspect** – concerning the practical application and use of ontologies

17

17

What is an Ontology?



18

18

Defining "Ontology"

- Ontologists generally have an intuitive understanding of what an ontology is.
- Expressing this understanding in precise words that withstand philosophical scrutiny is challenging.
- Consequently, there is no universally accepted definition of ontology at present.
- Over the past two decades, descriptions of ontology have gradually improved.

19

19

Defining "Ontology"

"An ontology is a specification of a conceptualization"

-Tom Gruber-

- This definition is considered unsatisfactory because it relies on vague terms:
  - Meaning of *"conceptualization"* and *"specification"* are ambiguous.
- This definition has been superseded by more precise alternatives.

20

20

Defining "Ontology"

"An ontology is described as a formal, explicit specification of a shared conceptualization."

- This definition is still unsatisfactory.
  - What exactly is meant by *conceptualization*?
  - What constitutes a *formal, explicit specification*?
  - To what extent must the ontology be *shared* for validity?

21

21

Defining "Ontology"

"An ontology is a logical theory accounting for the intended meaning of a formal vocabulary, i.e. its ontological commitment to a particular conceptualization of the world. The intended models of a logical language using such a vocabulary are constrained by its ontological commitment. An ontology indirectly reflects this commitment (and the underlying conceptualization) by approximating these intended models."

-Guarino-

22

22

Defining "Ontology"

"An ontology being equivalent to a Description Logic knowledge base."

-World Wide Web Consortium (W3C),

- In developing the OWL ontology language, W3C provided a simpler and more practical definition.

23

23

Defining "Ontology"

- In the Semantic Web context, ontologies are often regarded as equivalent to a logical theory, particularly a Description Logics knowledge base.
- Ontologists generally disapprove of referring to artefacts such as a thesaurus or an ER diagram represented in OWL as ontologies.

24

24

Defining "Ontology"

- Blurring of distinctions between different artefacts poses conceptual and practical problems.
- Representation in OWL alone does not necessarily make a construct an ontology.
- Conversely, constructs represented in languages other than OWL can still qualify as ontologies.

25

25

Ontologies: Good or Bad

- Ontologies, like software code, can be evaluated as good or bad.
- The assessment of ontologies is more elaborate than that of software code.
- In software, bad code may be unmaintainable, contain bugs, or fail to compile.
- For ontologies, a failure to comply with syntax is equivalent to code that does not compile.

26

26

Ontologies: Good or Bad

- Errors in ontologies can be of two types:
  - **Logical errors**: For example, a class cannot have any instances due to conflicting constraints.
  - **Semantic errors**: Ontology is logically correct but misrepresents the intended meaning (e.g., a class *Student* incorrectly classified as a subclass of *Table*).

27

27

Ontologies: Good or Bad

- Certain ontological structuring choices are constrained by ontological principles.
- Example: representing green apples can be approached in multiple ways:
  - As apples with the attribute "green."
  - As green objects with an apple shape.
- While logic treats both representations equivalently, intuitively the first approach is more reasonable.

28

28

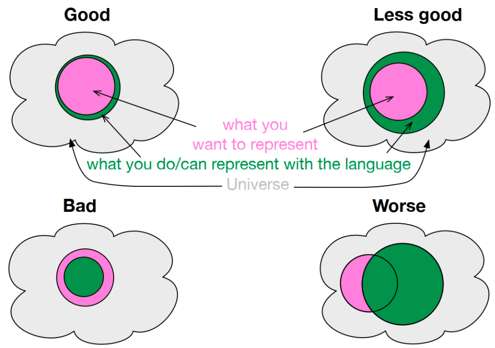
Ontologies: Good or Bad

- Reason:
  - "Apple" carries an identity condition and serves as a sortal, allowing objects to be individuated.
  - "Green" is merely a value of the attribute hasColor and does not confer identity.
- Ontology provides a framework for explaining such distinctions in knowledge representation.

29

29

Ontologies: Good or Bad



30

30

Ontologies: Good or Bad

- Quality of an ontology depends on the interplay between the:
  - Logic one uses to represent the knowledge in an ontology
  - Meaning of the entities in the subject domain
- Thus, ontologies can be classified as good or bad based on this interplay.

31

31

Ontologies: Good or Bad

- Precision and coverage are critical aspects that influence ontology quality.
- Both the choice of ontology language and the effectiveness of modelling practices affect precision and coverage.

32

32

Ontologies: Good or Bad

- **Good ontology:**
  - Intended knowledge is accurately represented.
  - Ontology closely matches the intended meaning with **high precision** and **maximum coverage**.
- **Less good ontology:**
  - Ontology includes excessive information beyond the intended scope.
  - Results in **low precision** but **maximum coverage**.

33

33

Ontologies: Good or Bad

Bad ontology:

- It has **high precision** but **limited coverage**.
- Essential concepts are missing, making the ontology inadequate for supporting the intended information system.

Worst ontology:

- Ontology has both **low precision** and **limited coverage**.
- It includes irrelevant information while omitting necessary concepts.

34

34

Uses of Ontologies

Data & Information System Integration

- A major motivation for developing an ontology is to have a shared vocabulary for data and information system integration.
  - Not just *data storage*
  - Not just *querying*
  - But **semantic alignment** across systems

35

35

Data & Information System Integration

- Example: Imagine two departments:
  - HR calls someone a **'Staff Member'**
  - Finance calls them an **'Employee'**
- Without alignment → miscommunication, errors, silos.
- Ontology provides a **common reference model** with formal meaning.

36

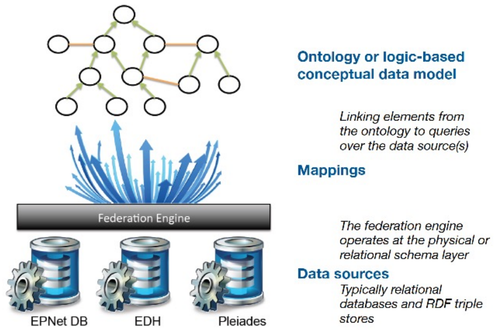
36

Ontology-Based Data Access (OBDA)

- **OBDA Architecture:**
  - **Ontology** – high-level conceptual model
  - **Mappings** – link ontology terms to database fields
  - **Query Rewriting** – SPARQL → SQL
- Result: Query the ontology, get data from legacy DBs

37

Ontology-Based Data Access (OBDA)



38

Ontologies as Part of a Solution

- Not standalone solutions
- But **key components** in larger systems: Natural Language Processing
- Scientific Discovery
- Question Answering
- Recommendation Systems

39

Case Study: Intelligent Question Answering

- **Scenario:** Chatbot for research papers
- **Input:** "What eats impalas in African savanna?"
- **Ontology Role:**

  Lion ⊑ eats.Impala

  Lion ⊑ Carnivore

  Inference: Lion → eats → Impala
- **Output:** "Lions eat impalas

40

Success Stories

- **Gene Ontology (GO)**
  - 40,000+ terms
  - Used in **>100,000 research papers**
  - Standard for gene function annotation
- **SNOMED CT**
  - 300,000+ clinical concepts
  - Used in **electronic health records worldwide**

41

Success Stories

- **GoodRelations**
  - E-commerce ontology
  - Acquired by Google (now in schema.org)
- **Ordnance Survey (UK)**
  - Spatial ontologies for mapping
  - Used in government, navigation
- **NASA SWEET Ontologies**
  - Earth science data integration
  - Enables climate model interoperability

42